Virology 377 (2008) 355–363

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

Four-tiered π interaction at the dimeric interface of HIV-1 integrase critical for DNA integration and viral infectivity

Laith Q. Al-Mawsawi [a], Anneleen Hombrouck [b], Raveendra Dayam [a], Zeger Debyser [b], Nouri Neamati [a,*]

[a] Department of Pharmacology and Pharmaceutical Sciences, School of Pharmacy, University of Southern California, Los Angeles, CA 90089, USA
[b] The Laboratory for Molecular Virology and Gene Therapy, KULeuven and IRC KULAK, Kapucijnenvoer 33, B-3000 Leuven, Flanders, Belgium

ARTICLE INFO

Article history:
Received 20 February 2008
Returned to author for revision
8 April 2008
Accepted 18 April 2008
Available online 2 June 2008

Keywords:
HIV-1 integrase
Strand transfer
π electron orbital interaction

ABSTRACT

HIV-1 integrase (IN) is an essential enzyme for viral infection. Here, we report an extensive π electron orbital interaction between four amino acids, W132, M178, F181 and F185, located at the dimeric interface of IN that is critical for the strand transfer activity alone. Catalysis of nine different mutant IN proteins at these positions were evaluated. Whereas the 3′-processing activity is predominantly strong, the strand transfer activity of each enzyme was completely dependent on an intact π electron orbital interaction at the dimeric interface. Four representative IN mutants were constructed in the context of the infectious NL4.3 HIV-1 viral clone. Whereas viruses with an intact π electron orbital interaction at the IN dimeric interface replicated comparable to wild type, viruses containing an abolished π interaction were non-infectious. Q-PCR analysis of viral DNA forms during viral replication revealed pleiotropic effects of most mutations. We hypothesize that the π interaction is a critical contact point for the assembly of functional IN multimeric complexes, and that IN multimerization is required for a functional pre-integration complex. The rational design of small molecule inhibitors targeting the disruption of this π–π interaction should lead to powerful anti-retroviral drugs.

Introduction

Viral cDNA integration within the host cell genome is essential for HIV replication. Once properly inserted, viral protein expression follows leading to maturation and propagation (Brown, 1997). Integration is facilitated by the viral protein integrase (IN) through two separate DNA reaction events. The first is 3′-processing, where a dinucleotide is hydrolytically cleaved off each viral DNA terminus. Cleavage results in two free 3′ hydroxyl groups used for a subsequent nucleophilic attack in the second reaction. 3′-processing occurs within the cytosol. Following this processing step, the IN-viral DNA complex, with a number of both viral and cellular proteins collectively referred to as the pre-integration complex (PIC), translocates to the nucleus. Here IN mediates the insertion of the processed viral DNA product within the host genome, which is termed strand transfer (Asante-Appiah and Skalka, 1999). IN reactions can be carried out in vitro using purified protein, a DNA substrate with ends mimicking the U3 or U5 viral DNA termini, and Mg^{2+} or Mn^{2+} as a cofactor (Bushman and Craigie, 1991).

Full length IN contains 288 amino acids and consists of three functional domains. The N-terminal domain (residues 1–50) contains a zinc binding motif consisting of two conserved pairs of histidine and cysteine residues (Johnson et al., 1986). The binding of zinc enhances IN multimerization (Lee et al., 1997; Zheng et al., 1996). The catalytic core domain (residues 51–212) is well conserved and contains three invariant acidic residues that make up the catalytic triad D,D(35)E motif that is essential for catalysis (Engelman and Craigie, 1992; Kulkosky et al., 1992; Leavitt et al., 1993). The C-terminal domain (residues 213–288) is much less evolutionarily conserved compared to the N-terminal and core domains, and possesses non-specific DNA binding activity (Khan et al., 1991; Vink et al., 1993; Woerner and Marcus-Sekura, 1993).

IN is known to function as a multimer. Early information provided by complementation studies showed that when N-terminal deletion IN proteins are mixed with C-terminal deletion IN mutants, which alone are non-functional, the catalytic activity of the protein is restored to near wild type (WT) levels (Engelman et al., 1993; van Gent et al., 1993), showing that IN functions at the very least as a dimeric species. Current models and experimental evidence on retroviral IN composition indicate that a dimeric species may be sufficient for 3′-processing, but a tetrameric arrangement, stabilized through DNA binding at the dimer–dimer interface, is necessary for the integration process and more relevant to the in vivo nucleo-protein complex (Bao et al., 2003; Karki et al., 2004). Correspondingly a recent fluorescence anisotropy study concluded that the HIV-1 IN dimer at the viral DNA extremities was the most proficient oligomeric species for 3′-processing, whereas a dimer of dimers was the predominant species for full integration, or strand transfer (Guiot et al., 2006).

Non-covalent chemical interactions involving aromatic rings are highly prevalent in intermolecular protein structures and contribute

* Corresponding author. Fax: +1 323 442 1390.
E-mail address: neamati@usc.edu (N. Neamati).

0042-6822/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2008.04.030

to the stability of monomeric protein folding and the proper arrangement of higher ordered oligomeric species. Aromatic ring interactions predominantly conform to three different orientations; a T-shaped edge-to-face orientation with an average distance of 4.96 Å, a parallel displaced stacking arrangement with an average distance of 3.5 Å, and finally a face-to-face stacking arrangement with an average distance of 3.5 Å. The interaction energy of aromatic–aromatic interaction is much weaker than hydrogen bonds (−4.78 kcal/mol), with the estimated stabilization energy of an aromatic–aromatic interaction in an α-helical secondary structure to be −0.8 kcal/mol (Meyer et al., 2003). The divalent sulfur atom of methionine residues also adopts a favored orientation when interacting with an aromatic face, and sulfhydryl–aromatic interactions have emerged as an authentic chemical interaction present in multiple protein examples. Sulfhydryl–aromatic interactions have an average distance of 3.7 Å (Pal and Chakrabarti, 2001), and have an estimated stabilization energy of −0.65 kcal/mol in an α-helical secondary structure (Waters, 2004). It is estimated that ~60% of aromatic residues (W, Y, and F) present in proteins are involved in π electron orbital interactions, with the T-shape edge-to-face and the parallel displaced stacking orientations being the most prevalent (Meyer et al., 2003). Although the energy contribution of the aromatic interaction is far weaker than the hydrogen bond in protein structure, these interactions have been recognized as playing an important role in overall protein stability.

The involvement of π electron aromatic interactions in a multitude of ligand–receptor contacts makes it increasingly important to understand these chemical interactions in more detail. Knowledge gained would better equip medicinal chemists in lead optimization and drug design targeting therapeutic protein sites containing aromatic residues found to be essential for enzymatic function. Previously, we identified an allosteric binding site for an HIV-1 IN inhibitor located at the dimeric interface (Al-Mawsawi et al., 2006). The presence of an aromatic interaction at the site was shown to be crucial for efficient inhibition of the enzyme by the small molecule. Here we have examined the π electron orbital interaction in more detail, and found it contains four different residues at the dimeric interface, and is absolutely indispensable for the strand transfer activity of the enzyme and is essential for viral replication. To our knowledge this is the first report of an internal non-covalent π electron interaction as being vital to the enzymatic function of a retroviral IN protein. This is also the first evidence identifying amino acid residues far removed from the IN active site that are critical for only one enzymatic function of the enzyme (strand transfer process), thus indirectly un-coupling the successive nature of the IN reaction. Previous studies have reported that IN proteins containing an Asp substitution at the Rous Sarcoma Virus IN position S124 (and the analogous S119 HIV-1 IN position) is severely defective for strand transfer catalysis, while increasing 3′-processing specificity (Konsavage et al., 2005; Konsavage et al., 2007). The serine position reported previously is directly adjacent to the DD(35)E catalytic triad. Here, we report a π electron orbital interaction at the IN dimeric interface, distant from the IN catalytic site, that is also required for strand transfer. Small molecules designed to disrupt this π electron orbital interaction should be effective allosteric binding IN inhibitors. We hypothesize that the non-covalent interaction identified here is a fundamental contact point in the assembly of fully functional HIV-1 IN multimers, which in turn is required for proper PIC formation and function.

Results

π electron orbital interaction at IN dimeric interface is required for strand transfer catalysis

Depicted in Fig. 1A is the four-tiered π electron orbital interaction at the dimeric interface. It involves residues M178, F181, and F185 from chain A, and residue W132 from chain B. The figure shown is a modified version of the cocrystal structure of the binding domains between LEDGF/p75 and IN (PDB 2B4J) (Cherepanov et al., 2005), with attention focused at the IN dimeric interface to emphasize the π

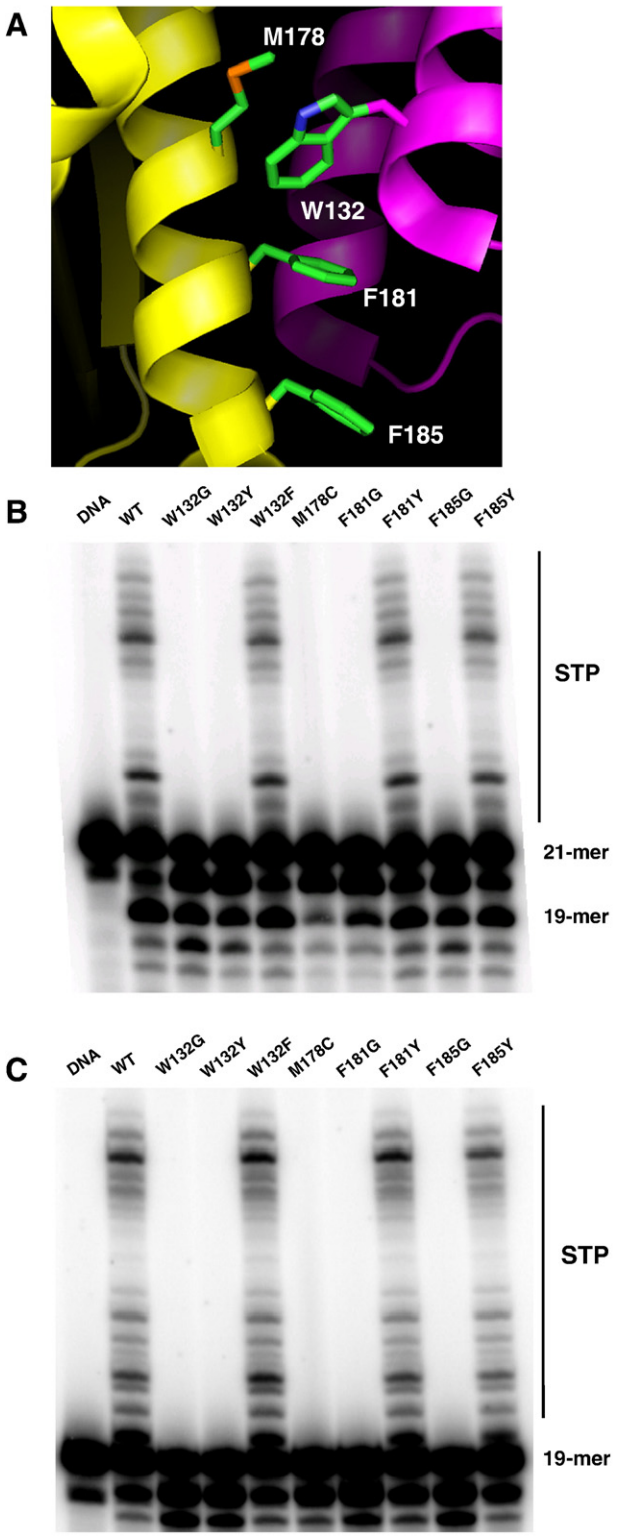

Fig. 1. The four-tiered aromatic interaction at the IN dimeric interface is critical for strand transfer catalysis. (A) Aromatic interaction between M178, W132, F181, and F185 at the dimeric interface (PDB 2B4J). Monomer A shown in yellow, monomer B in magenta; amino acid residues shown in stick representation, with atom colors: C (green), N (blue), and S (orange). (B) Relative activities of IN mutants using unprocessed 21-mer and (C) pre-processed 19-mer starting DNA substrate. STP refers to strand transfer products. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

interaction. Analysis of the non-covalent interaction reveals the sulfhydryl–aromatic distance between M178–W132 is 4.13 Å (shortest distance between S atom and ring center). The remaining interactions all appear to be variations of the T-shaped edge-to-face orientation upon visual inspection. The distance between the remaining interactions is 4.76 Å for W132–F181, and 5.65 Å for F181–F185 (shortest distance between edge–face). The interaction free enthalpy of the entire π electron orbital network is about −2.25 kcal/mol, or half the amount of a typical hydrogen bond, when treating the energy contribution of each interaction separately and assuming the affects are additive. It is important to note that all IN crystal structures containing the catalytic core domain we have examined have the non-covalent aromatic interaction at the dimeric interface in the same relative conformation, with the exception of F185. All existing IN crystal structures with the core domain present have F185 substituted for either lysine or histidine. IN proteins with the F185K or F185H substitution have been shown to be more soluble and therefore enhance the crystallization of the protein without affecting enzymatic activity (Jenkins et al., 1995). The crystal structure used here contained the F185K substitution. To generate Fig. 1A, we used the mutagenesis function in the molecular viewing program PYMOL (W. L. DeLano, www.pymol.org) to revert position 185 back to phenylalanine.

All mutant IN proteins used in this study were synthesized in the context of WT IN. Table 1 lists the activities of each mutant relative to WT, whereas a representative gel depicting the enzymatic activities of each protein is shown in Fig. 1B. Protein concentrations of each IN enzyme were determined prior to enzymatic assays to ensure equal loading and reliable functional comparisons. A clear specificity for a cohesive chemical interaction between these residues for the strand transfer process was observed. The substitutions generated at these four positions did not significantly affect IN 3′-processing activity, with the exception of M178C and F181G. The M178C mutant exhibited diminished 3′-processing activity (~30% of WT), whereas the protein containing a glycine substitution at position F181 showed a moderate decrease for this activity (~60% of WT). Interestingly, all the π electron orbital interaction mutants displayed an increase in 20-mer cleavage product (from 22% to 84%) as compared to WT. Three of the non-conservative substituted mutants additionally displayed an increase in 18-mer cleavage product (W132G: 88%, W132Y: 13%, and F185G: 12%) as compared to WT. This suggests residue substitutions in the π electron orbital interaction, both conservative and non-conservative, also influence the mechanics and/or regulation of 3′-processing. However, none of the IN mutants displayed a series of cleavage products below this size giving a clear indication that the phenomenon was not the result of bacterial exonuclease contamination that originated during recombinant protein purification. Additionally, the inhibitory effects of a LEDGF/p75 derived peptide displayed a propensity to inhibit both the classical 19-mer and aberrant 18-mer cleavage products, providing further evidence this is an IN-mediated 3′-processing irregularity (Al-Mawsawi et al., 2008). As for the strand transfer activity of the mutants, a strong dependence on the π electron orbital interaction was evidenced. For W132, the only residue involved in the interaction from chain B, a substitution to either glycine or tyrosine abolished the strand transfer catalysis of these mutant proteins. The only substitution to restore IN strand transfer activity at this position is phenylalanine. We also generated a lysine substitution at this position. W132K mutant proteins exhibited diminished 3′-processing capabilities with no strand transfer activity (data not shown). Based on the literature that sulfhydryl groups can interact favorably with aromatic rings, we included M178 in our analysis to determine if this residue could be involved in the proposed π aromatic interaction. Substituting a cysteine residue at this position created the most catalytically weak protein under study. As discussed, 3′-processing activity of this mutant was drastically reduced. The protein did not retain the ability to catalyze strand transfer. With regards to IN proteins containing substitutions at position F181, a non-conservative glycine substitution abolished strand transfer activity, whereas the conservative tyrosine substitution, where aromaticity is retained, restored this enzymatic function. Similarly, the mutant F185G did not retain strand transfer catalysis, whereas F185Y catalyzed this process efficiently. All mutants that retained strand transfer enzymatic function catalyzed this process comparably to WT.

IN reactions are successive in nature, with strand transfer activity dependent on prior 3′-processing catalysis. It is possible that our π electron orbital interaction deficient mutants, with the ability to catalyze 3′-processing but not strand transfer, could be deficient at a point immediately following 3′-processing before strand transfer catalysis. This would suggest that these mutants were 'locked' in a post-3′-processing catalytic conformation, and therefore unable to proceed towards strand transfer catalysis. To directly address at which catalytic point these mutants are deficient in the IN enzymatic process, we tested their function using a 3′-processed oligonucleotide substrate. A gel depicting the activity of each IN protein using this processed substrate is shown in Fig. 1C. The strand transfer catalytic profile of all IN proteins using this processed substrate was identical to the unprocessed DNA substrate. This result indicates that IN proteins containing a disrupted π aromatic interaction are deficient only for strand transfer activity. The π electron orbital interaction between residues W132, M178, F181, and F185 represents an ideal allosteric site for the design of strand transfer-specific small molecule IN inhibitors.

Table 1
Relative activities of integrase proteins

IN protein	3′-processing	Strand transfer
WT	+++[a]	+++
W132G	+++	–
W132Y	+++	–
W132F	+++	+++
M178C	+	–
F181G	++	–
F181Y	+++	+++
F185G	+++	–
F185Y	+++	+++

[a] Activity designation relative to wild type: –, 0–10%; +, 10–40%; ++, 40-80%;+++, 80-100%.

Aromaticity at position W132 is conserved among lentiviruses

Considering the proposed role of the π aromatic interaction for IN catalysis, we sought to determine the evolutionary importance of the four residues in a conservation sequence alignment. Eight lentiviral IN protein sequences were compared to the HIV-1 WT IN protein sequence. Also included in the alignment were three retroviral IN proteins of a different genus: human T-cell lymphotropic virus type 1 (*Deltaretrovirus*), Rous sarcoma virus (*Alpharetrovirus*), and Moloney murine leukemia virus (*Gammaretrovirus*). The results of the alignment are shown in Fig. 2. None of the four residues are absolutely conserved. At the position corresponding to W132 of HIV-1 however, aromaticity is completely conserved for lentiviral IN proteins.

Viruses containing an abolished π electron orbital interaction at the IN dimeric interface are non-infectious

To investigate the π interaction further, we constructed four representative IN mutants in the context of the infectious NL4.3 HIV-1 viral clone, i.e. NL4.3 M178C, F181G, F181Y and F185G. The mutant molecular clones were examined for their ability to replicate in MT-4 cells. Whereas the mutant virus containing an intact π electron orbital interaction at the IN dimeric interface (NL4.3 F181Y) replicated comparable to wild type, viruses containing an abolished π electron orbital interaction at the IN dimeric interface were non-infectious (NL4.3 M178C, F181G and F185G) (Fig. 3A).

To determine the step at which the replication cycle of these viruses is affected, real-time quantitative PCR analysis was performed

Fig. 2. Aromaticity at position W132 (HIV-1) is conserved in lentiviral IN proteins. In sequence alignment, highly conserved is depicted with black background with residues listed in white, while moderately conserved is shown in grey background with black lettering. '125' refers to starting position for HIV-1 IN protein sequence. W132 labeled with arrow; M178, F181, and F185 labeled with asterisk.

to quantify the formation of the different DNA species in the cell during infection. DNA was extracted at several time points post-infection. As expected, the replication-competent virus with a conservative tyrosine substitution at position F181 (that retains aromaticity) displayed similar levels of early total cDNA, 2-LTR circles and integrated proviruses as compared to WT HIV-1 (Figs. 3B,C,D). The three other mutants affected the accumulation of viral cDNA at early time points. The smallest effect was seen with the F185G mutant, whereas the strongest defect was observed for the M178C and the F181G mutants (Fig. 3B). All replication-incompetent viruses showed undetectable levels of 2-LTR circles and integrated proviral DNA in comparison to WT virus (Figs. 3C,D).

Model map of non-covalent interactions at the IN dimeric interface

Most protein–protein interfaces are characterized by smaller complementary pockets of either a hydrophobic nature or with buried hydrophilic charged residues forming salt bridges across the interface which are critical for the overall interaction (Moreira et al., 2007). To examine other possible interactions at the dimeric interface

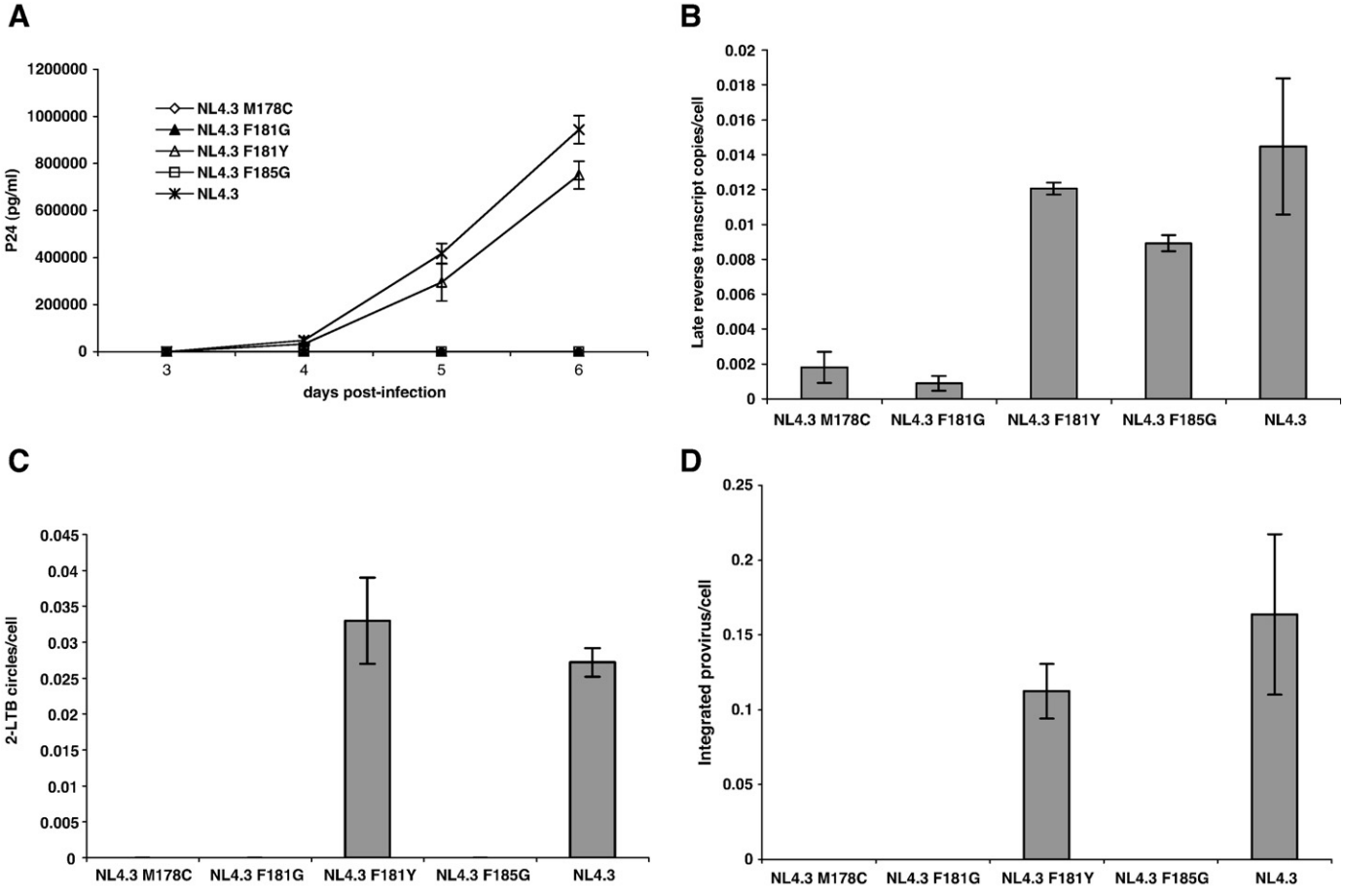

Fig. 3. Analysis of mutant HIV-1 virus replication. (A) MT-4 cells were inoculated with wild type or mutant viruses corresponding to 25 pg/mL p24: (*) WT NL4.3, (◊) NL4.3 M178C, (▲) NL4.3 F181G, (△) NL4.3 F181Y and (□) NL4.3 F185G. The replication kinetics were determined by measuring viral p24 antigen levels in the supernatant at different time points post-infection. Experiments were performed at least twice. Quantifications were performed in duplicate and averages±SD are shown. (B–C–D) Effects of mutations on viral DNA forms during replication. MT-4 cells were infected with WT or mutant HIV-1 at an MOI of 0.1. At different time points after infection, DNA extracts were prepared and analyzed by real-time PCR. Late reverse transcripts were quantified 6 h post-infection (B), and 2-LTR circles (C) and integrated proviral DNA at 72 h post-infection (D). Experiments were performed at least twice. Quantifications were performed in duplicate and averages±SD are shown.

of the catalytic core domain of IN, we have generated a model map of the catalytic core dimeric interface of IN using available crystal structures (Fig. 4A). Also depicted (Fig. 4B) is the protein sequence alignment discussed earlier spanning the other dimeric interface contacts we identified. The aromatic system between M178, W132, F181, and F185 is located at both horizontal extremes of the interface. We have also located another aromatic four-tiered interaction between W61 and W108, of chain A, which interacts with a corresponding W108 and W61 of chain B. The four tryptophan sidechains are each in the T-shaped edge-to-face orientation. The distance between each Trp is 4.54 Å for W61 (chain A)–W108 (chain A), 4.42 Å for W108 (chain A)–W108 (chain B), and 4.54 Å for W108 (chain B)–W61 (chain B) (shortest distance between edge–face). The interaction free enthalpy of the entire W61–W108 π electron orbital network to is about −2.4 kcal/mol, or half the amount of a typical hydrogen bond, when treating the energy contribution of each interaction separately and assuming the affects are additive. W61 is completely conserved in all the retroviral IN proteins included in our sequence alignment. A W61R substitution generated in both HIV-2 IN (van den Ent et al., 1998) and the Avian leukemia and sarcoma virus IN (Moreau et al., 2002) created a highly insoluble protein suggestive of defects in oligomerization. Indeed the aromatic interaction deficient mutants studied here were also more difficult to purify compared to IN proteins where the aromatic interaction was intact. At position W108, all the lentiviral IN proteins contain a residue capable of aromatic interaction. We have also located two highly charged symmetrical centers at the interface, one between E87 and K103, and the other between E85 and R107. Complimentary residues across an interface often co-evolve (Goh and Cohen, 2002), and salt bridge linkages are a common characteristic at protein–protein interfaces. When analyzing the sequence alignment these pairs of residues are highly correlated suggesting a distinct interaction. For each IN protein containing a glutamic acid residue at position 87 (corresponding to HIV-1), a positive residue (K or H) is located at the position corresponding to 103 of HIV-1 IN. A similar linkage can be observed for the E85 and R107 pair. It is difficult to determine how much stability these symmetrically placed pair of salt bridges may contribute to IN dimeric stability. A previous study examined the energetic contribution of two symmetrically placed Lys–Asp salt bridges using a homodimeric hemoglobin protein. It was determined that disruption of the charged pairs, provided by a Lys to Asp substitution, in combination with subsequent water hydration at the interface, disrupted dimeric stability by 8 kcal/mol (Ceci et al., 2002). Mutagenic studies conducted in HIV-2 IN found an E85W substitution did not affect function (van den Ent et al., 1998), however this substitution would create two closely-spaced cationic–π interactions, with each being energetically favorable by −0.4 kcal/mol (Waters, 2004). Conversely, both an E87Q and R107A substitution in HIV-2 IN were also reported as being catalytically active (van den Ent et al., 1998), indicating the symmetrical charged centers may only contribute a minor fraction to overall dimeric stability and are not as prominent a factor as the sequence alignment suggests.

Compilation and analysis of available alanine-scanning data on protein–protein interfaces show that 93% of residues that result in a change of overall free energy of binding higher than 4 kcal/mol

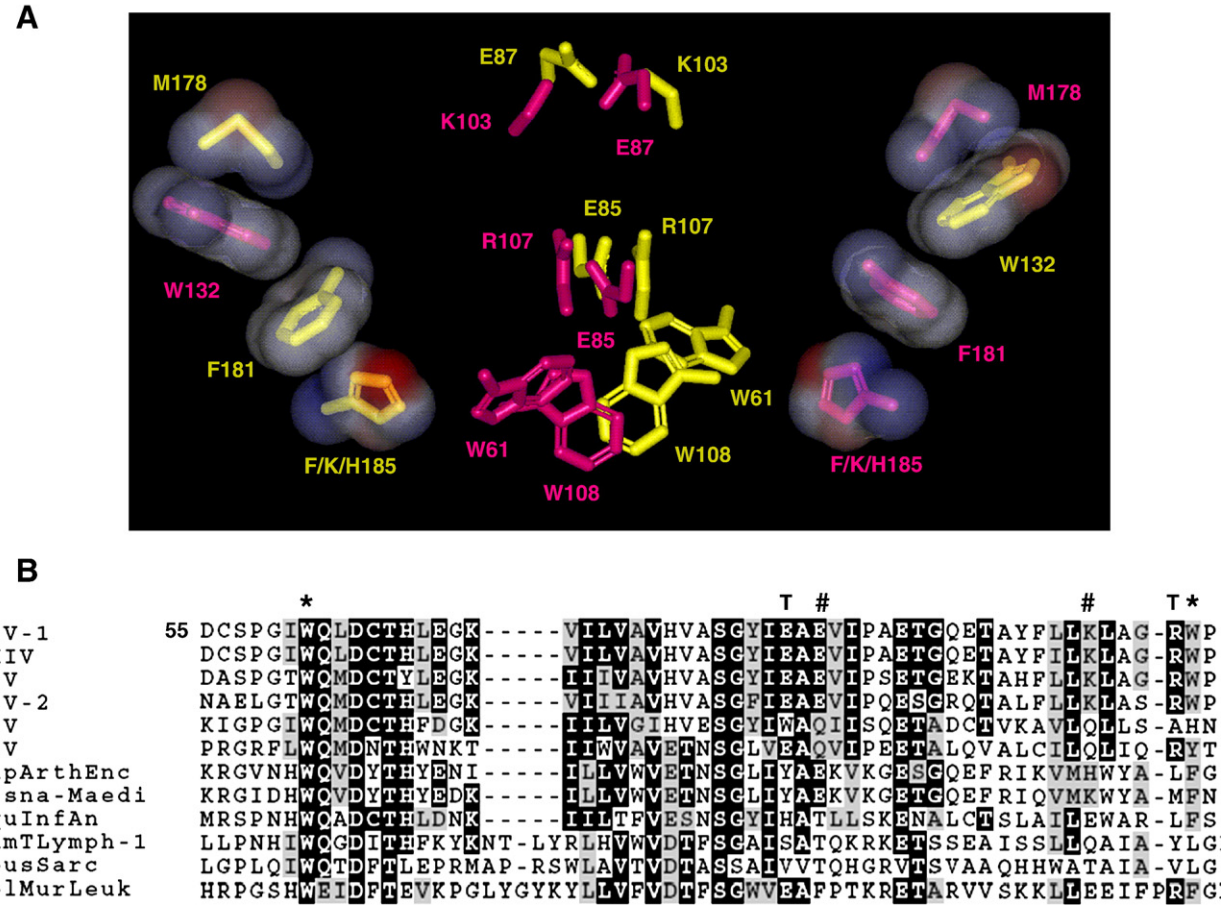

Fig. 4. IN catalytic core dimeric interface has four critical contact points, two aromatic and two charged. (A) Model map of the IN dimeric interface highlighting critical contact points between monomers. Residues from monomer A shown in yellow, monomer B in magenta; all residues shown in stick representation with electron orbital depiction for residues 132, 178, 181, and 185. Dimeric interface runs perpendicular to the plane of view. (B) Sequence alignment spanning IN region that contains other putative dimeric contact points; 55 refers to starting position for HIV-1 IN protein sequence. Amino acid pairs are denoted: W61 and W108 (*), E85 and R107 (T), E87 and K103 (#). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article)

following mutagenesis reside in complementary pockets (Keskin et al., 2005). If the energy of each π electron orbital interaction is considered as strong as the summation of bond strength, and each charged center is treated as a weak complementary pocket, it can be estimated that in combination these networks provide at least −10 kcal/mol in stabilization energy to the IN catalytic core domain dimer.

Discussion

We have located a π electron aromatic interaction between four residues at the dimeric interface of IN that is indispensable for strand transfer catalysis. Substitutions that disrupt this interaction abolished strand transfer activity with lesser effects on 3′-processing. The addition of other metal cofactors (Ca^{2+} and Zn^{2+}) in the reaction buffer does not restore strand transfer activity to the aromatic interaction deficient mutants as previously shown (Al-Mawsawi et al., 2006). W132 was the only residue from chain B involved in the aromatic interaction. Previously, we showed that W132 was critical for inhibition of a group of hydroxycoumarin IN inhibitors. Non-conservative substitutions (A, G, and R) decreased the potency of these molecules selectively, whereas a conservative tyrosine substitution, restoring the aromatic interaction, enabled the coumarins to bind correctly for optimal inhibition (Al-Mawsawi et al., 2006). All amino acid substitutions in this prior study were generated in the context of an IN double mutant (F185 K/C280S) that is known to promote solubility of the enzyme without affecting enzymatic activity in vitro (Jenkins et al., 1996). Here we generated all mutants in the context of WT IN. When we again generated the non-conservative glycine substitution at position W132 in the WT context, strand transfer activity was abolished as seen previously. Unexpectedly, when we generated the conservative W132Y substitution in the context of WT, strand transfer activity of the enzyme was also abolished. This prompted us to investigate the role of F185 within the aromatic interaction. It appears that the presence of F185K (soluble double mutant F185K/C280S) creates an environment at position W132 that is more tolerable, enabling a tyrosine to restore the aromatic interaction and therefore strand transfer catalysis. This is not entirely surprising considering single π aromatic interactions are often observed within larger cohesive π interactions within proteins (Meyer et al., 2003). The presence of a cationic element at one end of the aromatic interaction would have definite influence on the entire π aromatic system. In the context of WT IN, phenylalanine was the only substitution at W132 that could restore strand transfer activity to IN. W132, being the only residue from chain B in the interaction, seems to have a unique role in this aromatic system. The tryptophan residue at position 132 makes two contacts that bridge the dimeric interface. It is not presently clear if the sulfhydryl–aromatic interaction between M178–W132, as compared with the π edge-face interaction with W132–F181, play equal roles in this chemical system. W132 has also been implicated in one hotspot of interaction between the IN cellular cofactor LEDGF/p75, elucidated by the recent cocrystal structure and mutagenesis studies (Busschots et al., 2007; Cherepanov et al., 2005). Previously, an alanine substitution at W132 was shown to disrupt the LEDGF/p75-IN interaction (Busschots et al., 2007) and viruses harboring the IN W132A enzyme were replication defective due to a specific block in integration (Rahman et al., 2007). M178 was also observed in this contact point with LEDGF/p75. Making a conservative cysteine substitution at this position, where a sulfhydryl group was retained, created the most enzymatically compromised IN protein under study. It is possible that the addition of a hydrogen bond donor group (–SH) disrupted the sulfhydryl–aromatic interaction with W132. W132Y, with the phenyl linked hydrogen bond donor hydroxyl group, may have caused a similar sulfhydryl–aromatic disruption in the context of WT. To our knowledge, F181 has not been implicated in any IN process. Making a non-conservative glycine substitution at this position abolished strand transfer activity of the enzyme. Catalytic activity was restored with a conservative tyrosine substitution. Residue F185 followed a similar pattern as F181. A glycine substitution abolished strand transfer activity of the protein, whereas the mutant F185Y was fully active. Many studies have documented the strong enzymatic activity of IN proteins carrying F185K or F185H substitutions. It appears the π electron orbital interaction we have identified can tolerate either a histidine or lysine residue at position F185. Both these residues are highly capable of interacting in energetically favorable chemical aromatic systems (Burley and Petsko, 1985). Lysine and arginine residues have been well documented to make favorable cationic–π interactions with aromatic residues (Gallivan and Dougherty, 1999; Ma and Dougherty, 1997). Considering this leniency at position F185, we constructed W132K substitutions in both WT and the soluble double mutant (F185K/C280S). Both these proteins were severely hampered in 3′-processing activity and strand transfer catalysis was abolished (data not shown). Our observations indicate that substitution of cationic character can be tolerated at one end of the four-tiered aromatic system (F185), but not within the center of this interaction (W132).

To investigate the π interaction further, four IN substitutions (M178C, F181G, F181Y and F185G) were introduced in the infectious NL4.3 HIV-1 viral clone. Only the F181Y virus was replication-competent. All viruses containing an abolished π electron orbital interaction were non-infectious (Fig. 3A). Although these mutants mostly affect strand transfer activity in vitro, all virus mutants displayed pleiotropic effects in cellulo. M178C and F181G, but not F185G, significantly affected viral cDNA formation at 6 h post-infection (Fig. 3B). With the exception of F181Y, none of the IN-mutant-viral strains generated 2-LTR circles or integrated viral DNA to measurable levels. Reverse transcription and integration are coupled and partially synchronous during retroviral replication. In the pre-integration complex (PIC), IN binds to the first fragment of cDNA, well before the reverse transcription process is completed. A functional interaction between reverse transcriptase (RT) and IN has also been demonstrated (Wu et al., 1999), and C130S (residue adjacent to π interaction) IN-viral strains are defective in the initiation of endogenous reverse transcription (Zhu et al., 2004). We hypothesize that the π interaction is a critical interface for the assembly of fully functional multimeric complexes of IN, which further may be important for RT interactions and/or PIC stability. After 3′-processing, the PIC is transported to the nucleus, where the viral cDNA is integrated in the host DNA or non-productive 2-LTR circles are formed. 2-LTR circles are thus an indirect measure for nuclear import. The lack of nuclear import of our π interaction deficient IN mutants suggests that nuclear import of the PIC will only occur in the context of fully functional multimeric complexes of IN. Alternatively, erroneous 3′-processing by the mutant enzymes, as suggested by the appearance of alternative 20-mer cleavage products (Fig. 1B), may also interfere with the functionality of the PIC.

The association between the cellular cofactor LEDGF/p75 and IN is lentivirus-specific (Busschots et al., 2005). Here we have conducted a sequence alignment using 9 lentiviral IN protein sequences and the IN protein sequences from the three genuses observed not to interact with LEDGF/p75. At the position corresponding to W132 in HIV-1, aromaticity was completely conserved in lentiviral IN proteins, but not the IN proteins from alpha-, delta-, or gamma-retroviruses. The aromatic feature at this position may be a key element for lentiviral specificity of the LEDGF/p75-IN interaction. Also striking is that full length LEDGF/p75 protein binds to the area surrounding the four-tiered π electron orbital interaction and stimulates strand transfer activity, whereas our mutagenic study shows that the loss of this non-covalent interaction abolishes strand transfer activity. It has also been demonstrated that the overexpression of the IN binding domain of LEDGF/p75 inhibits viral replication (De Rijck et al., 2006) and this truncated version of LEDGF/p75 can not stimulate IN strand transfer catalysis like the full length protein (Cherepanov et al., 2004). The correlation between these phenomena is not presently clear, and may become apparent in future studies that incorporate substrate DNA in this nucleo-protein complex to give a more global mechanistic picture.

Recent studies suggest that an IN dimer is sufficient for 3′-processing, but a tetrameric arrangement, possibly stabilized by DNA substrate,

is required for strand transfer catalysis (Bao et al., 2003; Guiot et al., 2006; Karki et al., 2004). We have uncovered amino acid residues that are critical for strand transfer catalysis, but not 3′-processing. The π electron orbital network between M178, W132, F181, and F185, when intact, facilitates strand transfer catalysis. This suggests that the four-tiered π electron system is a critical pre-requisite for tetramerization of IN. It is possible that an intact aromatic system at this position helps stabilize the DNA substrate promoting IN tetramerization in order for strand transfer catalysis to proceed. Our results also indicate this process is required by RT function and/or PIC stability in the viral life cycle.

In conclusion, we have identified an extensive π electron orbital interaction at the dimeric interface of IN that involves both monomers, is four-tiered and includes residues M178, W132, F181, and F185. Viruses containing an abolished π electron orbital interaction at the IN dimeric interface were non-infectious. We hypothesize that the π interaction is a critical contact point for the assembly of fully functional multimeric complexes of IN, which is further required for strand transfer, RT function and/or PIC formation and function. Rationally designed small molecules targeting this four-tiered π electron system could provide powerful drugs for the treatment of HIV/AIDS.

Materials and methods

Site directed mutagenesis

Site directed mutagenesis was conducted on the pET-15b-IN plasmid, containing the WT sequence, a generous gift from Dr. Robert Craigie, Laboratory of Molecular Biology, NIDDK, NIH, Bethesda, MD. The plasmid contains full length IN fused to a 6-residue N-terminal histidine tag downstream from a T7 promoter. The Quickchange site-directed mutagenesis kit (Stratagene) was employed to make the point mutations according to the manufacturer's instructions. The oligonucleotide primers for each additional mutant are as follows, with the substituted nucleotide(s) shown in boldface and the mutated codon underlined. M178C: sense, 5′-GCT GAA CAT CTT AAG ACA GCA GTA CAA **TGC** GCA GTA TTC ATC CAC-3′; antisense, 5′-GTG GAT GAA TAC TGC **GCA** TTG TAC TGC TGT CTT AAG ATG TTC AGC-3′. W132F: sense, 5′-CC GCC TGT TGG T**TC** GCG GGG ATC AAG CAG G-3′; antisense, 5′-C CTG CTT GAT CCC CGC **GA**A CCA ACA GGC GG-3′. W132G: sense, 5′-GCC TGT TGG **G**GG GCG GGG ATC AAG CAG G-3′; antisense, 5′-C CTG CTT GAT CCC CGC C**C** CCA ACA GGC-3′. W132K: sense, 5′-GCC GCC TGT TGG **AA**G GCG GGG ATC AAG CAG G-3′; antisense, 5′-C CTG CTT GAT CCC CGC C**TT** CCA ACA GGC GGC-3′. W132Y: sense, 5′-GCC GCC TGT TGG T**AC** GCG GGG ATC AAG CAG G-3′; antisense, 5′-C CTG CTT GAT CCC CGC **GTA** CCA ACA GGC GGC-3′. F181G: sense, 5′-GCA GTA CAA ATG GCA GTA **GG**C ATC CAC AAT TTT AAA AGA AAA GGG-3′; antisense, 5′-CCC TTT TCT TTT AAA ATT GTG GAT G**CC** TAC TGC CAT TTG TAC TGC-3′. F181Y: sense, 5′-GCA GTA CAA ATG GCA GTA T**AC** ATC CAC AAT TTT AAA AGA AAA GGG-3′; antisense, 5′-CCC TTT TCT TTT AAA ATT GTG GAT G**TA** TAC TGC CAT TTG TAC TGC-3′. F185G: sense, 5′-G GCA GTA TTC ATC CAC AAT **GG**T AAA AGA AAA GGG GGG ATT G-3′; antisense, 5′-C AAT CCC CCC TTT TCT TTT A**CC** ATT GTG GAT GAA TAC TGC C-3′. F185Y: sense, 5′-G GCA GTA TTC ATC CAC AAT T**AT** AAA AGA AAA GGG GGG ATT G-3′; antisense, 5′-C AAT CCC CCC TTT TCT TTT A**TA** ATT GTG GAT GAA TAC TGC C-3′. Nucleotide replacements were confirmed by sequencing at the USC/Norris Comprehensive Cancer Center Microchemical Core Facility (University of Southern California).

Expression and purification of recombinant HIV-1 IN

The IN plasmid was expressed in *Escherichia coli* BL21(De3) PLYS S expression strain (Invitrogen) after induction by IPTG (1 mM) at an absorbance of 0.6–0.8 optical density at 595 nm. The culture was allowed to grow for an additional 3–4 h. at 37 °C. This was followed by centrifugation of the cells, at 3000 rpm in a bucket rotor centrifuge (Beckman) for 20 min. Pelleted cells were resuspended in lysis buffer (20 mM HEPES, pH 7.5, 5 mM imidazole, 100 mM NaCl) and passed twice through a French Press (Thermo Spectronic). Lysate was centrifuged at 31,000 *g* and the pellet was solubilized in a buffer containing 20 mM HEPES, pH 7.5, 5 mM imidazole, and 10 mM CHAPS. The IN protein was purified using Ni-affinity chromatography with Ni-NTA agarose beads (Qiagen), with increasing concentrations of imidazole used to elute the protein. Imidazole fractions were collected, and aliquots of each were separated on an SDS-PAGE gel and stained with Coomassie blue dye. Fractions containing IN protein were dialyzed in Spectra/Por molecular porous membrane tubing, MWCO 12–14,000 (Spectrum Laboratories, Inc.), suspended in a buffer containing 20 mM HEPES, pH 7.5, 500 mM NaCl, 40% glycerol, 0.2 mM EDTA, and 1 mM dithiothreitol (DTT). After dialysis, the purified enzyme solution contained 50 mM NaCl, 1 mM HEPES, pH 7.5, 50 μM EDTA, 50 μM DTT, and 10% glycerol (w/v).

Oligonucleotide substrates and IN enzymatic assays

Oligonucleotide substrates mimicking the HIV-1 U5 LTR DNA termini, 19-mer top strand (5′-GTGTGGAAAATCTCTAGCA-3′), 21-mer top strand (5′-GTGTGGAAAATCTCTAGCAGT-3′) and 21-mer bottom strand (5′-ACTGCTAGAGATTTTCCACAC-3′), were purchased from USC/Norris Comprehensive Cancer Center Microchemical Core Facility (University of Southern California) and purified by UV shadowing on polyacrylamide gel. In order to analyze the amount of both DNA reaction events, the 21-mer top oligonucleotide was 5′-end-labeled using T_4 polynucleotide kinase (Epicenter, Madison, WI) and γ-[^{32}P] ATP (MP Biomedicals, Irvine, CA). Afterwards the kinase was heat inactivated, and the 21-mer bottom oligonucleotide was added in 1.5 molar excess. The sample was then heated to 95 °C, cooled slowly, and subsequently passed through a G-25 Sephadex spin column (USA Scientific) to remove any unincorporated material from the annealed double-stranded oligonucleotide.

IN enzymatic assays for WT and each mutant followed the same general procedure. To determine the extent on both 3′-processing and strand transfer catalytic activities, purified IN enzyme (final concentration: 200 nM) was added to a reaction buffer [7.5 mM $MnCl_2$, 0.1 mg/mL bovine serum albumin, 10 mM β-mercaptoethanol, and 25 mM MOPS, pH 7.2] followed by the addition of 20 nM of the 5′-end ^{32}P-labeled linear double-stranded oligonucleotide, and the reaction mixture was incubated for 1 h at 30 °C. Upon completion the sample reactions were quenched by adding half the total volume (8 μL) of loading dye (98% deionized formamide, 10 mM EDTA, 0.025% xylene cyanol, and 0.025% bromophenol blue). A sample aliquot was electrophoresed on a denaturing 20% polyacrylamide gel (0.09 M Tris–borate, pH 8.3, 2 mM EDTA, 20% acrylamide, and 8 M urea). Gels were vacuum dried, exposed in a PhosphorImager cassette, analyzed using a Typhoon 8610 Variable Mode Imager (Amersham Biosciences), and quantified using ImageQuant 5.2.

Viral integrase alignment

The protein sequence for HIV-1 IN was obtained through direct translation of the wild type sequence expressed in the pET-15b-IN plasmid used in this study. All remaining integrase protein sequences from the *Retroviridae* family were obtained directly from the Retroviruses Homepage located at the National Center for Biotechnology Information (NCBI) database (http://www.ncbi.nlm.nih.gov/retroviruses/). The accession number for each species, following their respective genus is as follows: Lentivirus; SHIV (AAL78990), SIV (AAK97265), HIV-2 (AAC95341), FIV (NP040973), BIV (P19561), Caprine Arthritus Encephalitus Virus (P33459), Visna-Maedi Virus (AAM51650), and Equine Infectious Anemia Virus (NP056902). *Deltaretrovirus*; Human T-Cell Lymphotropic Virus-1 (P14708), *Alpharetrovirus*; Rous Sarcoma Virus (Q04095), *Gammaretrovirus*; Moloney Murine Leukemia

Virus (AAB03091). Protein sequences were subsequently imported into the multiple alignment software CLUSTAL W (Thompson et al., 1994) and exported in the boxshade results format.

Virus strains

The HIV-1 molecular clone pNL4.3 (Adachi et al., 1986) was obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH contributed by Dr. Malcolm Martin (Bethesda, MD). To generate the M178C, F181G, F181Y and F185G IN mutated viruses, site-directed mutagenesis was performed using the Kirsch and Joly method (Kirsch and Joly, 1998). The presence of the expected mutations was confirmed by DNA sequencing of the entire IN coding region. Virus productions were performed as described previously (De Rijck et al., 2006).

Determination of replication capacity of HIV-1 strains

Inoculants of various HIV-1 strains containing equal amounts of HIV-1 p24 antigen (25, 12.5 and 6.25 pg/mL) were added to MT-4 cells (50,000 cells/mL). From 3 days post-infection on, aliquots of cell free supernatants were harvested for the determination of viral p24 levels (Alliance HIV-1 P24 antigen ELISA Kit, Perkin Elmer Life Sciences, Milano, Italy).

HIV-1 infection assay

MT-4 cells (2.5×10^6 cells/well) were incubated with WT HIV-1 NL4.3 or mutant NL4.3 (corresponding to 250 ng of p24). After a 3 h incubation at 37 °C, the cells were washed with PBS, transferred to new medium and seeded in a 24-well plate (ca. 300,000 infected cells/well). In each 24-well plate, uninfected MT-4 cells were incubated in parallel. Each time a sample was prepared for Q-PCR analysis, an aliquot of uninfected cells was prepared as well.

Quantification of different HIV-1 DNA species by real-time PCR

A specifically designed set of TaqMan probe and primers was used to quantify the amount of each specific HIV-1 DNA form in the cell lysate. For the quantification of late reverse transcripts (total HIV-1 DNA), the forward primer 5′-TGTGTGCCCGTCTGTTGTGT-3′, the reverse primer 5′-GAGTCCTGCGTCGAGAGAGC-3′, and the probe 5′-(FAM)-CAGTGGCGCCCGAACAGGGA-(TAMRA)-3′ were used (Butler et al., 2001). For quantification of 2-LTR circles, the forward primer: 5′-GTGCCCGTCTGTTGTGTGACT-3′, the reverse primer: 5′-CTTGTCTTCTTTGGGAGAGAATTAGC-3′ and the probe: 5′-(FAM)-TCCA-CACTGACTAAAAGGGTCTGAGGGATCTCT-(TAMRA)-3′ were employed. Amplification was as reported previously (Van Maele et al., 2003). Quantification of integrated proviruses (Alu-PCR) was done using a nested Alu-PCR, modified from a method described previously (O'Doherty et al., 2002). The first round PCR was performed using the forward primer 5′-GCTAACTAGGGGAACCCACTGCTTA-3′ and the reverse primer, 5′-TGCTGGGATTACAGGCGTGAG-3′. Each sample was subjected to an initial denaturation of 5 min at 95 °C, followed by 15 amplification cycles of denaturation at 95 °C for 30 s, annealing at 60 °C for 40 s and extension at 72 °C for 1 min 30 s. The second round-run, real-time Taqman PCR was performed on 2.5 μL of the first round PCR product using the forward primer 5′-AGCTTGCCTTGAGTGCTTCAA-3′, the reverse primer 5′-TGACTAAAAGGGTCTGAGGGATCT-3′ and the probe 5′-FAM-TTACCAGAGTCACACAACAGACGGGCA-TAMRA-3′.

For each real-time PCR analysis, a standard curve was generated using dilutions of a representative standard of the amplicon being measured. No-template controls (no DNA added to the PCR mixture) and NACs (DNA extracted from cells that were not infected) were run with each experiment. In each sample RNAse P DNA was quantified as endogenous control and a standard curve for RNAse P was run in parallel. After initial incubations at 50 °C for 2 min and at 95 °C for 10 min, 40 cycles of amplification were carried out at 95 °C for 15 s, followed by 1 min at 60 °C. Reactions were analyzed by using the ABI Prism model 7700 sequence detection system (Applied Biosystems).

Acknowledgments

The following reagents were obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH: pINSD. His from Dr. Robert Craigie; pNL4-3 from Dr. Malcolm Martin and MT-4 from Dr. Douglas Richman. This work was supported by grants from The Campbell Foundation and the California HIV/AIDS Research Program (CHRP) to Nouri Neamati. Work at the K.U. Leuven was supported in part by the European Commission Targeting Replication and Integration of HIV (TRIoH) project LSHB-CT-2003-503480 and grants from the Fund for Scientific Research (FWO) (G.0497.05) and the Strategic Basic Research Program of the Institute for the Promotion of Innovation by Science and Technology in Flanders (IWT) (grant 530-030239).

References

Adachi, A., Gendelman, H.E., Koenig, S., Folks, T., Willey, R., Rabson, A., Martin, M.A., 1986. Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. J. Virol. 59 (2), 284–291.

Al-Mawsawi, L.Q., Fikkert, V., Dayam, R., Witvrouw, M., Burke Jr., T.R., Borchers, C.H., Neamati, N., 2006. Discovery of a small-molecule HIV-1 integrase inhibitor-binding site. Proc. Natl. Acad. Sci. U. S. A. 103 (26), 10080–10085.

Al-Mawsawi, L.Q., Christ, F., Dayam, R., Debyser, Z., Neamati, N., 2008. Inhibitory profile of a LEDGF/p75 peptide against HIV-1 integrase: insight into integrase–DNA complex formation and catalysis. FEBS Lett. 582 (10), 1425–1430.

Asante-Appiah, E., Skalka, A.M., 1999. HIV-1 integrase: structural organization, conformational changes, and catalysis. Adv. Virus Res. 52, 351–369.

Bao, K.K., Wang, H., Miller, J.K., Erie, D.A., Skalka, A.M., Wong, I., 2003. Functional oligomeric state of avian sarcoma virus integrase. J. Biol. Chem. 278 (2), 1323–1327.

Brown, P.O., 1997. Integration. In: Coffin, J.M., Hughes, S.H., Varmus, H.E. (Eds.), Retroviruses. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 161–203.

Burley, S.K., Petsko, G.A., 1985. Aromatic–aromatic interaction: a mechanism of protein structure stabilization. Science 229 (4708), 23–28.

Bushman, F.D., Craigie, R., 1991. Activities of human immunodeficiency virus (HIV) integration protein in vitro: specific cleavage and integration of HIV DNA. Proc. Natl. Acad. Sci. U. S. A. 88 (4), 1339–1343.

Busschots, K., Vercammen, J., Emiliani, S., Benarous, R., Engelborghs, Y., Christ, F., Debyser, Z., 2005. The interaction of LEDGF/p75 with integrase is lentivirus-specific and promotes DNA binding. J. Biol. Chem. 280 (18), 17841–17847.

Busschots, K., Voet, A., De Maeyer, M., Rain, J.C., Emiliani, S., Benarous, R., Desender, L., Debyser, Z., Christ, F., 2007. Identification of the LEDGF/p75 binding site in HIV-1 integrase. J. Mol. Biol. 365 (5), 1480–1492.

Butler, S.L., Hansen, M.S., Bushman, F.D., 2001. A quantitative assay for HIV DNA integration in vivo. Nat. Med. 7 (5), 631–634.

Ceci, P., Giangiacomo, L., Boffi, A., Chiancone, E., 2002. The mutation K30D disrupts the only salt bridge at the subunit interface of the homodimeric hemoglobin from *Scapharca inaequivalvis* and changes the mechanism of cooperativity. J. Biol. Chem. 277 (9), 6929–6933.

Cherepanov, P., Devroe, E., Silver, P.A., Engelman, A., 2004. Identification of an evolutionarily conserved domain in human lens epithelium-derived growth factor/transcriptional co-activator p75 (LEDGF/p75) that binds HIV-1 integrase. J. Biol. Chem. 279 (47), 48883–48892.

Cherepanov, P., Ambrosio, A.L., Rahman, S., Ellenberger, T., Engelman, A., 2005. Structural basis for the recognition between HIV-1 integrase and transcriptional coactivator p75. Proc. Natl. Acad. Sci. U. S. A. 102 (48), 17308–17313.

De Rijck, J., Vandekerckhove, L., Gijsbers, R., Hombrouck, A., Hendrix, J., Vercammen, J., Engelborghs, Y., Christ, F., Debyser, Z., 2006. Overexpression of the lens epithelium-derived growth factor/p75 integrase binding domain inhibits human immunodeficiency virus replication. J. Virol. 80 (23), 11498–11509.

Engelman, A., Craigie, R., 1992. Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function in vitro. J. Virol. 66 (11), 6361–6369.

Engelman, A., Bushman, F.D., Craigie, R., 1993. Identification of discrete functional domains of HIV-1 integrase and their organization within an active multimeric complex. EMBO J. 12 (8), 3269–3275.

Gallivan, J.P., Dougherty, D.A., 1999. Cation–pi interactions in structural biology. Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9459–9464.

Goh, C.S., Cohen, F.E., 2002. Co-evolutionary analysis reveals insights into protein–protein interactions. J. Mol. Biol. 324 (1), 177–192.

Guiot, E., Carayon, K., Delelis, O., Simon, F., Tauc, P., Zubin, E., Gottikh, M., Mouscadet, J.F., Brochon, J.C., Deprez, E., 2006. Relationship between the oligomeric status of HIV-1 integrase on DNA and enzymatic activity. J. Biol. Chem. 281 (32), 22707–22719.

Jenkins, T.M., Hickman, A.B., Dyda, F., Ghirlando, R., Davies, D.R., Craigie, R., 1995. Catalytic domain of human immunodeficiency virus type 1 integrase: identification

of a soluble mutant by systematic replacement of hydrophobic residues. Proc. Natl. Acad. Sci. U. S. A. 92 (13), 6057–6061.
Jenkins, T.M., Engelman, A., Ghirlando, R., Craigie, R., 1996. A soluble active mutant of HIV-1 integrase: involvement of both the core and carboxyl-terminal domains in multimerization. J. Biol. Chem. 271 (13), 7712–7718.
Johnson, M.S., McClure, M.A., Feng, D.F., Gray, J., Doolittle, R.F., 1986. Computer analysis of retroviral pol genes: assignment of enzymatic functions to specific sequences and homologies with nonviral enzymes. Proc. Natl. Acad. Sci. U. S. A. 83 (20), 7648–7652.
Karki, R.G., Tang, Y., Burke Jr., T.R., Nicklaus, M.C., 2004. Model of full-length HIV-1 integrase complexed with viral DNA as template for anti-HIV drug design. J. Comput. Aided Mol. Des. 18 (12), 739–760.
Keskin, O., Ma, B., Nussinov, R., 2005. Hot regions in protein–protein interactions: the organization and contribution of structurally conserved hot spot residues. J. Mol. Biol. 345 (5), 1281–1294.
Khan, E., Mack, J.P., Katz, R.A., Kulkosky, J., Skalka, A.M., 1991. Retroviral integrase domains: DNA binding and the recognition of LTR sequences. Nucleic Acids Res. 19 (4), 851–860.
Kirsch, R.D., Joly, E., 1998. An improved PCR-mutagenesis strategy for two-site mutagenesis or sequence swapping between related genes. Nucleic Acids Res. 26 (7), 1848–1850.
Konsavage Jr., W.M., Burkholder, S., Sudol, M., Harper, A.L., Katzman, M., 2005. A substitution in Rous sarcoma virus integrase that separates its two biologically relevant enzymatic activities. J. Virol. 79 (8), 4691–4699.
Konsavage Jr., W.M., Sudol, M., Lee, N.E., Katzman, M., 2007. Retroviral integrases that are improved for processing but impaired for joining. Virus Res. 125 (2), 198–210.
Kulkosky, J., Jones, K.S., Katz, R.A., Mack, J.P., Skalka, A.M., 1992. Residues critical for retroviral integrative recombination in a region that is highly conserved among retroviral/retrotransposon integrases and bacterial insertion sequence transposases. Mol. Cell. Biol. 12 (5), 2331–2338.
Leavitt, A.D., Shiue, L., Varmus, H.E., 1993. Site-directed mutagenesis of HIV-1 integrase demonstrates differential effects on integrase functions in vitro. J. Biol. Chem. 268 (3), 2113–2119.
Lee, S.P., Xiao, J., Knutson, J.R., Lewis, M.S., Han, M.K., 1997. Zn2+ promotes the self-association of human immunodeficiency virus type-1 integrase in vitro. Biochemistry 36 (1), 173–180.
Ma, J.C., Dougherty, D.A., 1997. The cationminus signpi interaction. Chem. Rev. 97 (5), 1303–1324.
Meyer, E.A., Castellano, R.K., Diederich, F., 2003. Interactions with aromatic rings in chemical and biological recognition. Angew. Chem. Int., Ed. Engl. 42 (11), 1210–1250.
Moreau, K., Faure, C., Verdier, G., Ronfort, C., 2002. Analysis of conserved and non-conserved amino acids critical for ALSV (avian leukemia and sarcoma viruses) integrase functions in vitro. Arch. Virol. 147 (9), 1761–1778.
Moreira, I.S., Fernandes, P.A., Ramos, M.J., 2007. Hot spots—a review of the protein–protein interface determinant amino-acid residues. Proteins 68 (4), 803–812.
O'Doherty, U., Swiggard, W.J., Jeyakumar, D., McGain, D., Malim, M.H., 2002. A sensitive, quantitative assay for human immunodeficiency virus type 1 integration. J. Virol. 76 (21), 10942–10950.
Pal, D., Chakrabarti, P., 2001. Non-hydrogen bond interactions involving the methionine sulfur atom. J. Biomol. Struct. Dyn. 19 (1), 115–128.
Rahman, S., Lu, R., Vandegraaff, N., Cherepanov, P., Engelman, A., 2007. Structure-based mutagenesis of the integrase-LEDGF/p75 interface uncouples a strict correlation between in vitro protein binding and HIV-1 fitness. Virology 357 (1), 79–90.
Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22 (22), 4673–4680.
van den Ent, F.M., Vos, A., Plasterk, R.H., 1998. Mutational scan of the human immunodeficiency virus type 2 integrase protein. J. Virol. 72 (5), 3916–3924.
van Gent, D.C., Vink, C., Groeneger, A.A., Plasterk, R.H., 1993. Complementation between HIV integrase proteins mutated in different domains. EMBO J. 12 (8), 3261–3267.
Van Maele, B., De Rijck, J., De Clercq, E., Debyser, Z., 2003. Impact of the central polypurine tract on the kinetics of human immunodeficiency virus type 1 vector transduction. J. Virol. 77 (8), 4685–4694.
Vink, C., Oude Groeneger, A.M., Plasterk, R.H., 1993. Identification of the catalytic and DNA-binding region of the human immunodeficiency virus type I integrase protein. Nucleic Acids Res. 21 (6), 1419–1425.
Waters, M.L., 2004. Aromatic interactions in peptides: impact on structure and function. Biopolymers 76 (5), 435–445.
Woerner, A.M., Marcus-Sekura, C.J., 1993. Characterization of a DNA binding domain in the C-terminus of HIV-1 integrase by deletion mutagenesis. Nucleic Acids Res. 21 (15), 3507–3511.
Wu, X., Liu, H., Xiao, H., Conway, J.A., Hehl, E., Kalpana, G.V., Prasad, V., Kappes, J.C., 1999. Human immunodeficiency virus type 1 integrase protein promotes reverse transcription through specific interactions with the nucleoprotein reverse transcription complex. J. Virol. 73 (3), 2126–2135.
Zheng, R., Jenkins, T.M., Craigie, R., 1996. Zinc folds the N-terminal domain of HIV-1 integrase, promotes multimerization, and enhances catalytic activity. Proc. Natl. Acad. Sci. U. S. A. 93 (24), 13659–13664.
Zhu, K., Dobard, C., Chow, S.A., 2004. Requirement for integrase during reverse transcription of human immunodeficiency virus type 1 and the effect of cysteine mutations of integrase on its interactions with reverse transcriptase. J. Virol. 78 (10), 5045–5055.